A GENERALISATION OF THE DELOGNE-KÅSA METHOD FOR FITTING HYPERSPHERES

Emanuel E. Zelniker and I. Vaughan L. Clarkson

School of Information Technology & Electrical Engineering
The University of Queensland
Queensland, 4072, AUSTRALIA
{zelniker,v.clarkson}@itee.uq.edu.au

ABSTRACT

In this paper, we examine the problem of fitting a hypersphere to a set of noisy measurements of points on its surface. Our work generalises an estimator of DELOGNE (*Proc. IMEKO-Symp. Microwave Measurements* 1972, 117-123) which he proposed for circles and which has been shown by Kåsa (*IEEE Trans. Instrum. Meas.* 25, 1976, 8-14) to be convenient for its ease of analysis and computation. We also generalise CHAN's 'circular functional relationship' to describe the distribution of points. We derive the CRAMÉR-RAO lower bound (CRLB) under this model and we derive approximations for the mean and variance for fixed sample sizes when the noise variance is small. We perform a statistical analysis of the estimate of the hypersphere's centre. We examine the existence of the mean and variance of the estimator for fixed sample sizes. We find that the mean exists when the number of sample points is greater than $M+1$, where M is the dimension of the hypersphere. The variance exists when the number of sample points is greater than $M+2$. We find that the bias approaches zero as the noise variance diminishes and that the variance approaches the CRLB. We provide simulation results to support our findings.

1. INTRODUCTION

In this paper, we examine the problem of fitting a hypersphere to a set of noisy measurements of points on the hypersphere's surface. Historically, research has tended to focus on the 2-dimensional problem, namely, fitting circles. The accurate fitting of a circle to noisy measurements of points on its circumference is an important and much-studied problem in the scientific literature. Applications of this problem include archæology [1], geodesy [2], physics [3, 4] microwave engineering [5] and computer vision and metrology [6]. The estimation of a circle's centre and its radius appears to have been first studied by THOM [1], who proposes an approximate method of least squares to fit circles to ancient stone rings in megalithic sites in Britain and Scotland. A complete formulation of the solution to this problem for circles and spheres by the method of least squares is given by ROBINSON [2].

SPÄTH [7] provided a new descent algorithm for circle fitting, which he later generalised to spheres [8]. He considers an objective function of the maximum likelihood estimator (MLE) which minimises the radial errors between the given data points and the true sphere. A numerical algorithm for the MLE is provided which uses two types of iterating steps, by partitioning the set of parameters. The algorithm is initiated from a point which is evaluated using KÅSA's algorithm [5] which uses a least-squares fit. The problem is monotonically convergent, but only to local minima. The convergence of this algorithm as well as the algorithm in [6] are discussed in [9].

There are several statistical models which describe the positions of the noisy circle points on the circumference of a circular arc. The appropriate choice of model is of course heavily dependent on the application. The first detailed statistical analysis of any model to be published appears to be that of CHAN [10]. He proposes a 'circular functional relationship', which assumes that the measurement errors are instances of independent and identically distributed (i.i.d) random variables and, furthermore, the points are assumed to lie at fixed but unknown angles around the circumference. In other words, not only the centre and radius of the circle are unknown parameters to be estimated, but so are the angles of each circumferential point. CHAN derives a method to find the MLE when the errors have a Gaussian distribution. This method is identical to the least-squares method of [2]. He also examines the consistency of the estimator. Other models which describe the positions of the noisy circle points are discussed and examined by BERMAN & CULPIN [11].

BERMAN & CULPIN [11] have carried out a statistical analysis of the MLE and the DELOGNE-KÅSA estimator (DKE) for circles. Specifically, they prove some results regarding the asymptotic consistency and variance of the estimates. ZELNIKER & CLARKSON [12, 13] examine the properties of the DKE for fixed (small) sample sizes rather than its asymptotic properties. They show that the DKE centre estimates have moments under certain conditions. Specifically, the expectation of the DKE exists if the number of sample points is greater than 3 and variance exists when this number is greater than 4. They also show that, although the DKE is known to be biased and asymptotically inefficient, as the noise variance approaches zero, the bias approaches zero and the variance approaches the CRAMÉR-RAO lower bound (CRLB).

In this paper, our aim is to adopt the framework developed in [12] and extend the fixed-sample-size statistical analysis of the DKE for circle points to the generalised DKE (GDKE) where the set of noisy measurements of points lie on a hypersphere's surface of arbitrary dimension. An important feature in this paper is the derivation of the CRLB for a hypersphere's centre estimate. CHAN & THOMAS [14] and KANATANI [15] derive the CRLB for circles, but we extend the derivation to hyperspheres and to our knowledge, this has not been done before. We set out condi-

Emanuel Zelniker is additionally supported by a scholarship from the Commonwealth Scientific & Industrial Research Organisation.

tions for which the mean and variance of the GDKE for a hypersphere's centre exist for fixed sample sizes under the hyperspherical model, which is a generalisation of CHAN's circular functional model with Gaussian errors. Where the mean exists, we show that the estimator is unbiased in the limit as noise variance approaches zero. Where the variance exists, we show that the variance approaches the CRLB as the noise variance approaches zero. We rely on stochastic matrix theory in our analysis, in particular, the WISHART distribution. We provide simulation results to support our findings.

2. STATISTICAL MODEL FOR A HYPERSPHERE

CHAN's circular functional model [10] assumes N measured points on the circumference of a 2-dimensional circle. The measurement process introduces random errors so that the 2-dimensional Cartesian coordinates $\mathbf{p}_i = (x_i, y_i)$, $i = 1, \ldots, N$ can be expressed as

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} r\cos\theta_i \\ r\sin\theta_i \end{pmatrix} + \begin{pmatrix} \xi_i \\ \eta_i \end{pmatrix},$$

where (a, b) is the centre of the circle, r is its radius, the θ_i are the angles around the circumference on which the points lie and the ξ_i and η_i are instances of random variables representing the measurement error. They are assumed to be zero-mean and i.i.d and are Gaussian with variance σ^2.

Extending this model to a hypersphere is relatively straightforward. We now have a point in an M-dimensional space represented by $\mathbf{y}_j = (y_{1,j}, \ldots, y_{M,j})^T$, $i = 1, \ldots, M$, $j = 1, \ldots, N$ such that

$$\mathbf{y}_j = \mathbf{s}_j + \boldsymbol{\xi}_j. \tag{1}$$

Here, $\mathbf{s}_j = \mathbf{c} + r\mathbf{u}_j$ where $\mathbf{c} = (c_1, \ldots, c_M)^T$ is the centre of the hypersphere, r is its radius, the $\mathbf{u}_j$ are unit vectors and $\boldsymbol{\xi}_j = (\xi_{1,j}, \ldots, \xi_{M,j})^T$ are instances of random variables representing the measurement error. They are assumed to be zero-mean i.i.d. In addition, we will specify that they are Gaussian with variance σ^2. In this paper, we explicitly exclude the possibility that $r = 0$ or that $\mathbf{u}_1 = \mathbf{u}_2 = \ldots = \mathbf{u}_N$.

Figure 1 shows some data with N points for the surface of a hypersphere, $\mathbf{y}_1, \ldots, \mathbf{y}_N$, displaced from the surface by noise.

3. THE CRAMÉR-RAO LOWER BOUND FOR HYPERSPHERES

We will now derive the CRLB for a hypersphere's centre.

Theorem 1. *For an unbiased estimator of a hypersphere's centre* $\mathbf{c}$*, according to the statistical model (1), the variances of the estimator* $\mathbf{c}$ *can not be less than the diagonal elements of the inverse of*

$$\frac{1}{r^2\sigma^2}(\mathbf{S}^T\mathbf{P}\mathbf{S}), \tag{2}$$

where $\mathbf{S} = (\mathbf{s}_1 \cdots \mathbf{s}_N)^T$ *and the matrix* $\mathbf{P}$ *is an* $N \times N$ *idempotent matrix defined so that* $\mathbf{P} = \mathbf{I} - (\mathbf{1}\mathbf{1}^T/N)$ *where* $\mathbf{I}$ *is the* $N \times N$ *identity matrix and* $\mathbf{1}$ *is an* N*-dimensional column vector, all of whose entries are* 1*. Note that* $\|\mathbf{P}\|_2 = 1$.

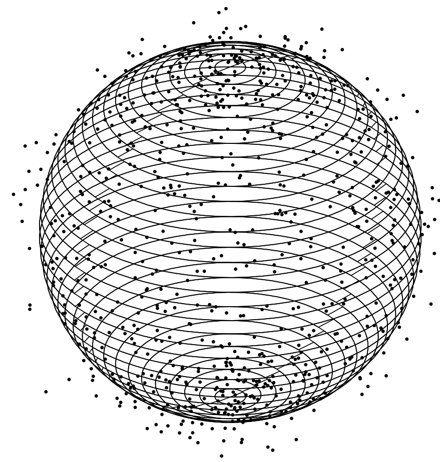


Fig. 1. An example of noisy measurements of points on the surface of a hypersphere.

Outline of proof. The logarithm of the conditional probability density function for $\mathbf{y}_1, \ldots, \mathbf{y}_N$ is as follows

$$L(\mathbf{y}_1, \ldots, \mathbf{y}_N \mid \boldsymbol{\Omega}) = -MN\log(\sqrt{2\pi\sigma^2}) - \frac{1}{2\sigma^2}\sum_{j=1}^{N}\|\boldsymbol{\xi}_j\|_2, \tag{3}$$

where $\boldsymbol{\Omega} = (\mathbf{c}, r, \theta_{k,j})$ and $\theta_{k,j}$, $k = 1, \ldots, M-1$ are spherical coordinates. That is, we express $\mathbf{u}_j$ so that

$$\mathbf{u}_j = \begin{pmatrix} \cos\theta_{M-1,j}\prod_{k=1}^{M-2}\cos\theta_{k,j} \\ \sin\theta_{M-1,j}\prod_{k=1}^{M-2}\cos\theta_{k,j} \\ \sin\theta_{M-2,j}\prod_{k=1}^{M-3}\cos\theta_{k,j} \\ \sin\theta_{M-3,j}\prod_{k=1}^{M-4}\cos\theta_{k,j} \\ \vdots \\ \sin\theta_{2,j}\cos\theta_{1,j} \\ \sin\theta_{1,j} \end{pmatrix}.$$

From VAN TREES [16, pp. 79-80], for an unbiased estimator of a hypersphere's centre $\mathbf{c}$, radius r and angles $\theta_{k,j}$, the variances of the estimators of $\mathbf{c}$, r, $\theta_{k,j}$ can not be less than the diagonal elements of the inverse of the FISHER Information Matrix, $\mathbf{J}$, where

$$\mathbf{J} \triangleq E\left[\frac{\partial L}{\partial \boldsymbol{\Omega}_p}\frac{\partial L}{\partial \boldsymbol{\Omega}_q}\right] = -E\left[\frac{\partial^2 L}{\partial \Omega_p \partial \Omega_q}\right]. \tag{4}$$

Therefore, we need to partially differentiate (3) with respect to the c_i, r and $\theta_{k,j}$ in order to construct (4). The parameter vector $\boldsymbol{\Omega}$ and (4) tells us that $\mathbf{J}$ is an $[(M+1) + N(M-1)] \times [(M+1) +$

$N(M-1)]$ symmetric matrix. We find that

$$\frac{\partial L}{\partial c_i} = \frac{1}{\sigma^2}\sum_{j=1}^{N}\xi_{i,j}, \tag{5}$$

$$\frac{\partial L}{\partial r} = \frac{1}{\sigma^2}\sum_{j=1}^{N}\sum_{i=1}^{M}\xi_{i,j}u_{i,j}, \tag{6}$$

$$\frac{\partial L}{\partial \theta_{k,j}} = -\frac{r}{\sigma^2}\sum_{j=1}^{N}\sum_{t=1}^{M-k}\xi_{t,j}u_{t,j}\tan\theta_{t,j} - \xi_{M-k+1,j}\prod_{u=1}^{k}\cos\theta_{u,j}. \tag{7}$$

Now, because we are interested in the CRLBs of $\mathbf{c}$ only, we can partition the FISHER Information Matrix (4) as

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{pmatrix}.$$

Setting the matrix $\mathbf{J}_{11}$ to be the upper $M \times M$ sub-matrix of $\mathbf{J}$ and substituting (5), (6) and (7) into (4) followed by the use of the block matrix inversion lemma (see [17]), the CRLBs of $\mathbf{c}$ will lie along the diagonal of the upper $M \times M$ sub-matrix of $\mathbf{J}^{-1}$, which is the inverse of (2). □

4. MAXIMUM LIKELIHOOD ESTIMATION

If we define $r_j(\mathbf{c}) = \|\mathbf{y}_j - \mathbf{c}\|_2$ where $\|\cdot\|_2$ represents the Euclidean norm, then it can be shown that the MLE is

$$(\hat{\mathbf{c}}_{\text{ML}}, \hat{r}_{\text{ML}}) = \arg\min_{(\mathbf{c},r)} \sum_{j=1}^{N}[r_j(\mathbf{c}) - r]^2. \tag{8}$$

The difficulties with the MLE are that it is hard to analyse and also to compute numerically. Analytically, it is not certain a priori that a global minimum exists, or whether there might be local minima [11, 9]. Numerically, the only methods available for solution are iterative. This raises the usual issues with convergence and sensitivity to the initial solution estimate.

5. THE GENERALISED DELOGNE-KÅSA ESTIMATOR

The analytical and numerical difficulties with the MLE in Section 4 as applied to circles led KÅSA [5] to propose the use of a modified estimator, originally due to DELOGNE [18], which we call the DKE. The generalisation of the DKE, or GDKE, can be written as

$$(\hat{\mathbf{c}}_{\text{GDK}}, \hat{r}_{\text{GDK}}) = \arg\min_{(\mathbf{c},r)} \sum_{j=1}^{N}[r_j^2(\mathbf{c}) - r^2]^2. \tag{9}$$

Notice that by squaring the argument of the sum there are no expressions involving *square roots* as is otherwise implied by the use of the Euclidean norm in (8). The linearisation which results from this formulation simplifies the analysis and the computation considerably. It can be shown that this estimator is a standard linear least-squares estimator. In terms of matrix algebra, we find that

$$\hat{\mathbf{c}}_{\text{GDK}} = \tfrac{1}{2}(\mathbf{PY})^{\#}\mathbf{P}f(\mathbf{Y}). \tag{10}$$

Here, the superscript '#' represents the MOORE-PENROSE generalised inverse or pseudo-inverse and, for a matrix $\mathbf{A}$ where $\mathbf{A}^T\mathbf{A}$ is non-singular, we may write $\mathbf{A}^{\#} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$. Further,

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1^T \\ \vdots \\ \mathbf{y}_N^T \end{pmatrix},\ E[\mathbf{Y}] = \mathbf{S} \text{ and } f(\mathbf{Y}) = \begin{pmatrix} \|\mathbf{y}_1\|_2^2 \\ \vdots \\ \|\mathbf{y}_N\|_2^2 \end{pmatrix}. \tag{11}$$

6. ANALYSIS OF THE GDKE

We now turn our attention to the analysis of the GDKE for fixed sample sizes. We are firstly interested in the question of whether the mean and variance exist. Then, we state low-variance approximations for their values which are valid whenever they exist.

Before outlining the proof for the main theorem in this section, we observe the following lemmas. The full proofs of all theorems and lemmas are generalisations of those to be found in [12].

Lemma 1. *The matrix $\mathbf{P}$ has a singular-value decomposition of the form*

$$\mathbf{P} = \boldsymbol{\Upsilon}\boldsymbol{\Delta}\boldsymbol{\Upsilon}^T$$

where $\boldsymbol{\Upsilon}$ is an orthogonal matrix and $\boldsymbol{\Delta} = \operatorname{diag}\{1,\ldots,1,0\}$.

Lemma 2. *For any vectors $\mathbf{x}, \boldsymbol{\mu} \in \mathbb{R}^N$,*

$$\exp\left(-\frac{\|\mathbf{x}-\boldsymbol{\mu}\|_2^2}{2\sigma^2}\right) \le \exp\left(\frac{\|\boldsymbol{\mu}\|_2^2}{2\sigma^2}\right)\exp\left(-\frac{\|\mathbf{x}\|_2^2}{4\sigma^2}\right).$$

Corollary 1. *If $\mathbf{X} = (X_1,\ldots,X_N)^T$ is a multivariate normal random vector such that each $X_i \sim N(\mu_i,\sigma^2)$ is independent, then*

$$E[\|f(\mathbf{X})\|_2^k] \le 2^{N/2}\exp\left(\frac{\|\boldsymbol{\mu}\|_2^2}{2\sigma^2}\right)E[\|f(\mathbf{Y})\|_2^k],$$

where $\boldsymbol{\mu} = (\mu_1,\ldots,\mu_N)^T$ and $\mathbf{Y} = (Y_1,\ldots,Y_N)^T$ is a multivariate normal random vector such that each $Y_i \sim N(0,2\sigma^2)$ is independent.

Definition 1. *We say that an $N \times n$ matrix $\mathbf{X}$ is a rectangular Gaussian matrix if each element is i.i.d. with identical variance σ^2 and $E[\mathbf{X}] = \boldsymbol{\mu}$. We denote its distribution $G(N,n,\boldsymbol{\mu},\sigma^2)$.*

Theorem 2. *The mean of the GDKE for a hypersphere's centre, as defined in (10), exists if the number of sample points on the surface, N, is greater than $M+1$, where M is the dimension of the hypersphere.*

Outline of proof. If the variance σ^2 is zero then $\hat{\mathbf{c}}_{\text{GDK}}$ is deterministic. In this case, the mean clearly exists, since the pseudo-inverse of $\mathbf{PY}$ always exists. Hence, we restrict our attention to the case where $\sigma^2 > 0$.

In order to show that the expectation exists, it is necessary to show that $E[\|\hat{\mathbf{c}}_{\text{GDK}}\|_2] < \infty$. Notice that $\mathbf{Y}$ has distribution $G(N,M,\mathbf{S},\sigma^2)$ and from the definition of expectation and the sub-multiplicative inequality

$$E[\|\hat{\mathbf{c}}_{\text{GDK}}\|_2] \le \tfrac{1}{2}E[\|(\mathbf{PY})^{\#}\|_2\|\mathbf{P}f(\mathbf{Y})\|_2]. \tag{12}$$

Using the notation of Lemma 1, we can define $\boldsymbol{\Upsilon}^T\mathbf{Y}$ as follows

$$\boldsymbol{\Upsilon}^T\mathbf{Y} = \begin{pmatrix} \mathbf{F} \\ \bar{\mathbf{y}} \end{pmatrix},$$

where $\mathbf{F} \sim G(N-1, M, \boldsymbol{\mu}_F, \sigma^2)$ and $\boldsymbol{\mu}_F$ is all but the last row of $\boldsymbol{\Upsilon}^T\mathbf{S}$ and $\bar{\mathbf{y}} \sim G(1, M, \boldsymbol{\mu}_{\bar{y}}, \sigma^2)$ with $\boldsymbol{\mu}_{\bar{y}}$ the last row of $\boldsymbol{\Upsilon}^T\mathbf{S}$. Notice that $\boldsymbol{\Upsilon}^T\mathbf{Y} \sim G(N, M, \boldsymbol{\Upsilon}\mathbf{S}, \sigma^2)$. Then

$$\boldsymbol{\Delta}\boldsymbol{\Upsilon}^T\mathbf{Y} = \begin{pmatrix}\mathbf{F}\\ \mathbf{0}\end{pmatrix}.$$

Also, using Lemma 1,

$$\|(\mathbf{P}\mathbf{Y})^{\#}\|_2 = \|\mathbf{F}^{\#}\|_2. \tag{13}$$

Now, since $\|\mathbf{P}f(\mathbf{Y})\|_2 \leq \|\mathbf{P}\|_2\|f(\mathbf{Y})\|_2 = \|f(\mathbf{Y})\|_2$, we can say that (with $\|\cdot\|_F$ representing the FROBENIUS norm of its argument)

$$\begin{aligned}\|f(\mathbf{Y})\|_2 &\leq \|f(\mathbf{Y})\|_1 = \|\mathbf{Y}\|_F^2\\ &\leq N\|\mathbf{Y}\|_2^2 = N\|\boldsymbol{\Upsilon}^T\mathbf{Y}\|_2^2 \leq N\|\boldsymbol{\Upsilon}^T\mathbf{Y}\|_F^2\\ &= N(\|\mathbf{F}\|_F^2 + \|\bar{\mathbf{y}}\|_F^2)\\ &\leq N^2(\|\mathbf{F}\|_2^2 + \|\bar{\mathbf{y}}\|_2^2).\end{aligned} \tag{14}$$

We have therefore bounded the expression for $\|\mathbf{P}f(\mathbf{Y})\|_2$ above by a polynomial in terms of $\|\mathbf{F}\|_2$ and $\|\bar{\mathbf{y}}\|_2$ which we denote as $p_1(\|\mathbf{F}\|_2, \|\bar{\mathbf{y}}\|_2)$. Thus, we can now say that

$$E[\|\hat{\mathbf{c}}_{\text{GDK}}\|_2] \leq \tfrac{1}{2}E[\|\mathbf{F}^{\#}\|_2 p_1(\|\mathbf{F}\|_2, \|\bar{\mathbf{y}}\|_2)]. \tag{15}$$

Also, notice that $\mathbf{F}$ and $\bar{\mathbf{y}}$ are independent. Therefore, it is clear that we can take the expectation with respect to $\bar{\mathbf{y}}$ to show that

$$E[\|\hat{\mathbf{c}}_{\text{GDK}}\|_2] \leq \tfrac{1}{2}E[\|\mathbf{F}^{\#}\|_2 p_2(\|\mathbf{F}\|_2)],$$

where $p_2(\|\mathbf{F}\|_2)$ is a polynomial in $\|\mathbf{F}\|_2$ only. Through the use of Corollary 1, we find that

$$E[\|\hat{\mathbf{c}}_{\text{GDK}}\|_2] \leq 2^{(N/2)-1}\exp\left(\frac{\|\boldsymbol{\mu}_F\|_F^2}{2\sigma^2}\right)E[\|\mathbf{W}^{\#}\|_2 p_2(\|\mathbf{W}\|_2)], \tag{16}$$

and $\mathbf{W}$ is a random matrix like $\mathbf{F}$ but each element has zero mean and twice the variance, *i.e.* $\mathbf{W} \sim G(N-1, M, \mathbf{0}, 2\sigma^2)$.

Consider the value of $\|\mathbf{W}\|_2$ and $\|\mathbf{W}^{\#}\|_2$, *i.e.*,

$$\begin{aligned}\|\mathbf{W}\|_2 &= v_1, \\ \|\mathbf{W}^{\#}\|_2 &= \frac{1}{v_M},\end{aligned} \tag{17}$$

where v_1 and v_M are the maximum and minimum singular values of $\mathbf{W}$ respectively. Therefore, they are the square roots of the maximum and minimum eigenvalues of $\mathbf{W}^T\mathbf{W}$ which has a WISHART distribution.

From MUIRHEAD [19, p. 106], the exact joint density function for the n eigenvalues of a general WISHART matrix can be written as

$$P_{\Lambda_1,\ldots,\Lambda_n}(\lambda_1,\ldots,\lambda_n) = \begin{cases} K_{N,n}\exp\left(-\dfrac{1}{2}\displaystyle\sum_{i=1}^n \lambda_i\right)\displaystyle\prod_{i=1}^n \lambda_i^{(N-n-1)/2}\prod_{i<q}(\lambda_i - \lambda_q) & \text{if } \lambda_1 \geq \cdots \geq \lambda_n, \\ 0 & \text{otherwise,}\end{cases} \tag{18}$$

where $K_{N,n}$ is a normalising constant and where the random matrix giving rise to the Wishart matrix is $G(N, n, \mathbf{0}, 1)$.

In our case, the random matrix is $G(N-1, M, \mathbf{0}, 2\sigma^2)$, and writing the density function (18) in terms of the singular values $v_1, \ldots, v_M$, we have

$$P_{v_1,\ldots,v_M}(v_1,\ldots,v_M) = \begin{cases} \dfrac{4K_{N-1,M}}{(2\sigma^2)^{N-1}}\exp\left(-\dfrac{1}{4\sigma^2}\displaystyle\sum_{i=1}^M v_i^2\right)\displaystyle\prod_{i=1}^M v_i^{N-M-1}\prod_{i<q}(v_i^2 - v_q^2) & \text{if } v_1 \geq \cdots \geq v_M, \\ 0 & \text{otherwise.}\end{cases} \tag{19}$$

Looking at (16) and using (17), we can see that

$$E[\|\mathbf{W}^{\#}\|_2 p_2(\|\mathbf{W}\|_2)] = \int_0^\infty\int_{v_1}^\infty\cdots\int_{v_{M-1}}^\infty \frac{p_2(v_1)}{v_M} P_{v_1,\ldots,v_M}(v_1,\ldots,v_M)\,dv_1\cdots dv_M, \tag{20}$$

It can now be seen from (20) that we have bounded $E[\|\hat{\mathbf{c}}_{\text{GDK}}\|_2]$ above by an M-dimensional integral in $v_1, \ldots, v_M$. This integral is the product of a degree-2 polynomial of non-negative powers of $v_1, \ldots, v_M$ with an exponential of the negative square of $v_1, \ldots, v_M$ when $N \geq M + 2$. Such an integral is finite, *e.g.*, see [20, §3.461]. □

For the remainder of this section, we omit the proofs of our theoretical results. Each uses a variation on the proof method of Theorem 2, similar to the development for circles in [12].

Theorem 3. *The variance of the GDKE for a hypersphere's centre, as defined in (10), exists if the number of sample points on the surface, N, is greater than $M + 2$.*

Theorem 4. *When the mean of the GDKE exists,*

$$E[\hat{\mathbf{c}}_{GDK}] = \mathbf{c}^T + O(\sigma). \tag{21}$$

Theorem 5. *When the variance of the GDKE exists,*

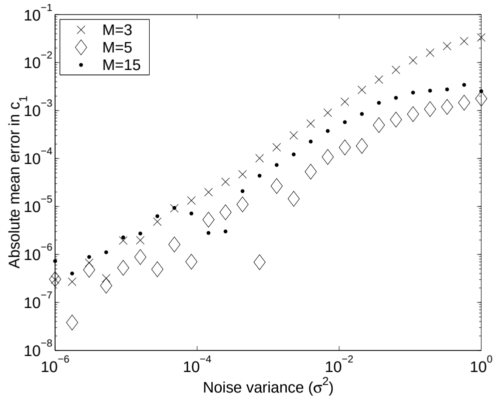
$$\operatorname{var}(\hat{\mathbf{c}}_{GDK}, \hat{\mathbf{c}}_{GDK}) = r^2\sigma^2(\mathbf{S}^T\mathbf{P}\mathbf{S})^{-1} + O(\sigma^3). \tag{22}$$

Note that $r^2\sigma^2(\mathbf{S}^T\mathbf{P}\mathbf{S})^{-1}$ is the upper $M \times M$ sub-matrix of $\mathbf{J}^{-1}$, where $\mathbf{J}$ is the FISHER Information Matrix defined in (4).

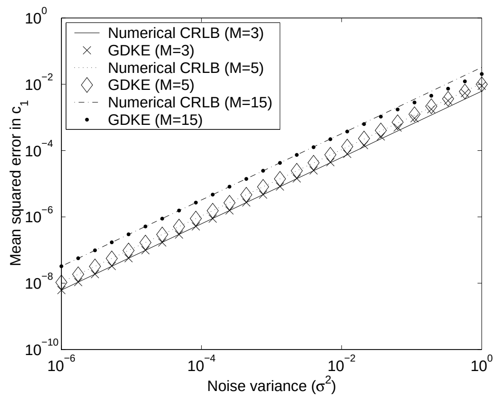
7. RESULTS FROM SIMULATION

Simulations were performed for hyperspheres of dimension 3, 5 and 15, *i.e.* $M = 3, 5, 15$. For each M, the GDKE was simulated using a Monte-Carlo analysis. 500 arbitrary points with no noise were generated with a uniform distribution around the full hypersphere's surface. The radius r was set to 1. Equation (2) was used to generate the CRLB by constructing the matrix in (2), taking its inverse and considering the values along the main diagonal of the inverse. In each trial, noise in the form of $\boldsymbol{\xi}_j$ was added to the true points and the GDKE was run repeatedly, $100\,000$ times, for each value of σ to obtain estimates for the centre of the hypersphere $\hat{\mathbf{c}}$ and use them to generate mean error values and mean square error (MSE) values. The amount of noise, σ was varied from 10^{-3} to 1 in equal geometric increments.

The absolute mean error values in $\hat{c}_1$ are plotted versus σ^2 in Figure 2(a) on a logarithmic scale for $M = 3, 5, 15$. It can be seen that the mean error decreases with decreasing σ. This



(a)



(b)

Fig. 2. Simulation results for variying σ for a full hypersphere of Dimension $M = 3, 5, 15$.

is consistent with Theorems 2 and 4. The MSE values in $\hat{c}_1$ are plotted against their corresponding CRLB for the same level of noise σ in Figure 2(b) on a logarithmic scale. It can be seen that at high values of noise, the estimator $\hat{c}_1$ departs from the CRLB. However, as the noise level, σ approaches zero, the estimator $\hat{c}_1$ approaches the CRLB. This is consistent with Theorems 3 and 5.

We have chosen to illustrate the results with plots for $\hat{c}_1$. Plots for $\hat{c}_2, \ldots, \hat{c}_M$ where $M = 3, 5, 15$ follow an identical pattern.

8. CONCLUSION

In Theorem 2, we showed that the expectation of the GDKE exists if $N > M + 1$ and, in Theorem 3, we showed that the variance of the GDKE exists if $N > M + 2$. In the limit as the noise variance approaches zero, the estimates have been shown to be unbiased in Theorem 4 and to be statistically efficient in Theorem 5. The results from simulation demonstrate that the GDKE quickly approaches the CRLB as noise variance is reduced.

9. REFERENCES

[1] A. Thom, "A Statistical Examination of the Megalithic Sites in Britain," *J. Roy. Statist. Soc. Ser. A General*, vol. 118, pp. 275–295, Mar. 1955.

[2] Stephen M. Robinson, "Fitting Spheres by the Method of Least Squares," in *Comm. ACM*, 1961, p. 491.

[3] J. F. Crawford, "A non-iterative method for fitting circular arcs to measured points," in *Nucl. Instru. Methods Phys. Research*, 1983, pp. 223–225.

[4] V. Karimäki, "Effective circle fitting for particle trajectories," in *Nucl. Instru. Methods Phys. Research*, 1991, pp. 187–191.

[5] I. Kåsa, "A circle fitting procedure and its error analysis," *IEEE Trans. Instrum. Meas.*, vol. 25, pp. 8–14, Mar. 1976.

[6] U. M. Landau, "Esimation of a Circular Arc Center and its Radius," *Comput. Vision Graph. Image Process.*, vol. 38, pp. 317–326, June 1986.

[7] H. Späth, "Least-Square Fitting by Cirles," *Computing*, vol. 57, pp. 179–185, 1996.

[8] H. Späth, "Least-Square Fitting with Spheres," *J. Optim. Theory Appl.*, vol. 96, pp. 191–199, Jan. 1998.

[9] N. Chernov and C. Lesort, "Least squares fitting of circles and lines," to appear in Journal of Mathematical Imaging and Vision.

[10] Nai Ng Chan, "On Circular Functional Relationships," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 27, pp. 45–56, Apr. 1965.

[11] Mark Berman and David Culpin, "The Statistical Behaviour of Some Least Squares Estimators of the Centre and Radius of a Circle," *J. Roy. Statist. Soc. Ser. B Stat. Methodol.*, vol. 48, pp. 183–196, Oct. 1986.

[12] E. E. Zelniker and I. Vaughan L. Clarkson, "A Statistical Analysis of the DELOGNE-KÅSA Method for Fitting Circles," submitted to Digital Signal Processing, unpublished.

[13] Emanuel E. Zelniker and I. Vaughan L. Clarkson, "A Statistical Analysis of Least Squares Circle-Centre Estimation," in *IEEE Internat. Symp. Signal Process. Inform. Tech.*, Darmstadt, Germany, Dec. 2003.

[14] Y. T. Chan and Samuel M. Thomas, "Cramér-Rao Lower Bound for Estimation of a Circular Arc Center and its Radius," in *Graph. Models Image Process.*, Nov. 1995, vol. 57, pp. 527–532.

[15] K. Kanatani, "Cramér-Rao Lower Bounds for Curve Fitting," in *Graph. Models Image Process.*, 1998, pp. 93–99.

[16] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*, Wiley, New York, 1968.

[17] D. V. Ouellette, "Schur Complements and Statistics," *Lin. Alg. Appl.*, vol. 36, pp. 187–295, Mar. 1981.

[18] P. Delogne, "Computer Optimization of Deschamps' Method and Error Cancellation in Reflectometry," in *Proc. IMEKO-Symp. Microwave Measurements*, Budapest, Hungary, 1972, pp. 117–123.

[19] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

[20] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals Series and Products*, Academic Press, San Diego, 2000.